



4845141

Лившиц Евгений Давидович

## Сходимость жадных алгоритмов

01.01.01 — вещественный, комплексный и функциональный  
анализ

А в т о р е ф е р а т  
диссертации на соискание ученой степени  
доктора физико-математических наук

Москва — 2011

12 МАЙ 2011

Работа выполнена на кафедре нелинейного анализа и оптимизации факультета физико-математических и естественных наук Российского университета дружбы народов.

**Официальные оппоненты:**

доктор физико-математических наук, профессор Колягин С.В.  
доктор физико-математических наук, профессор Лукашенко Т.П.  
доктор физико-математических наук, доцент Терехин П.А.

**Ведущая организация:**


Институт Математики и Механики Уральского отделения РАН.

Защита диссертации состоится 17 мая 2011 г. в 16 ч. 00 мин. на заседании диссертационного совета Д.212.203.27 в Российском университете дружбы народов по адресу: 115419, Москва, ул. Орджоникидзе, д. 3, ауд. 495а.

С диссертацией можно ознакомиться в Учебно-научном информационном библиотечном центре (Научной библиотеке Российского университета дружбы народов) по адресу 117198, Москва, ул. Миклухо-Маклая д. 6.

Автореферат разослан “ \_\_\_\_ ” \_\_\_\_\_ 2011 г.

Ученый секретарь диссертационного совета  
кандидат физико-математических наук,  
доцент

 Россовский Л.Е.

## Общая характеристика работы.

**Актуальность темы.** Теория приближения относится к тем областям математики, которые тесно связаны с прикладными задачами естествознания и техники. Увеличение мощности вычислительных систем, происшедшее в последние десятилетия, позволило приступить к решению новых более вычислительноемких задач. Одной из них является построение "индивидуального приближения" для заданной функции  $f$  из некоторого класса  $K$ . Приближение предлагается строить как элемент линейного подпространства  $L$  из заранее определенного (по  $K$ ) семейства линейных подпространств  $\mathcal{L}$ . При этом выбор  $L \in \mathcal{L}$  зависит от  $f$ , что отличает эту задачу от классической задачи аппроксимации класса  $K$ , где приближающее подпространство  $L$  выбиралось единым для всех  $f \in K$ . Другими словами, класс  $K$  приближается нелинейным объектом  $\bigcup_{L \in \mathcal{L}} L$ . Такие приближения имеют также важное теоретическое значение. Как было показано Р.С. Исмагиловым [2], К.И. Осколковым [6] и В.Е. Майоровым [5] этот нелинейный (индивидуальный) подход может давать преимущества в некоторых классических задачах. Изучение оценок снизу для этого метода приближения было начато В.С. Кашиным [3].

Начиная с 80-х годов 20-го века, в рамках математической статистики и теории приближения стали интенсивно изучаться *жадные алгоритмы*, что, с одной стороны, дало теоретическое обоснование ряда методов вычислительной математики, а, с другой стороны, позволило получить конструктивные способы нахождения наилучших  $m$ -членных приближений. Ключевую роль в формировании теории жадных алгоритмов сыграли работы Дж. Фридмана, В. Стузла, С. Маллата, Ж. Жанга, П. Хьюбера, Л. Джоунса, А. Бэррона, Р. ДеВора, В.Н. Темлякова, С.В. Коныгина, Д. Донохо и др. В наши дни жадные алгоритмы успешно применяются в задачах распознавания образов, медицины, финансовой математики, обработки сигналов и ряде других областей.

Необходимо отметить, что со временем большая часть результатов по жадным алгоритмам начала формулироваться и доказываться на языке теории функций. Более того, идеология теории функций в значительной степени определила направления дальнейших исследований жадных алгоритмов.

Пусть  $X = (X, \|\cdot\|)$  — действительное банахово пространство. Множество  $\mathcal{D}$ ,  $\mathcal{D} \subset X$ , будем называть словарем, если оно состоит из векто-

105

ров с единичной нормой, и замыкание линейной оболочки  $\mathcal{D}$  совпадает со всем  $X$ :

$$g \in \mathcal{D} \Rightarrow \|g\| = 1, \overline{\text{span}}\mathcal{D} = X.$$

Для произвольного элемента  $f \in X$  и  $m \in \mathbb{N}$  требуется найти конечную линейную комбинацию  $m$  элементов словаря, достаточно хорошо приближающую  $f$ :

$$f \rightarrow \sum_{k=1}^m c_k(f) g_k(f), \quad c_k(f) \in \mathbb{R}, \quad g_k(f) \in \mathcal{D}. \quad (1)$$

Особенностями данной постановки являются:

- от  $f$  зависят не только коэффициенты  $c_k(f)$ , но и элементы словаря  $g_k(f)$ , участвующие в приближении,
- словарь  $\mathcal{D}$  может быть как базисом, так и переполненной системой. Важными примерами переполненных систем являются словарь плоских волн (ridge-функций), RBF-словари, переполненные системы случайных векторов и т.д.

При этом желательно, чтобы величина ошибки аппроксимации  $\|f - \sum_{k=1}^m c_k(f) g_k(f)\|$  была близка к значению наилучшего  $m$ -членного приближения (это понятие было введено С.Б. Стечкиным в [9])

$$\sigma_m(f) := \sigma_m(f, \mathcal{D}) := \sigma_m(f; \mathcal{D}, X) = \inf_{c_1, \dots, c_m \in \mathbb{R}, g_1, \dots, g_m \in \mathcal{D}} \|f - \sum_{k=1}^m c_k g_k\|.$$

В большинстве случаев, особенно для переполненных словарей, построение  $m$ -членных приближений (1), пусть даже не наилучших, является весьма важной и интересной задачей.

Наметим подход к построению  $m$ -членных приближений с помощью жадных алгоритмов.

Пусть  $X$  является сепарабельным гильбертовым пространством  $H = (H, \langle \cdot, \cdot \rangle)$  с нормой  $\|\cdot\| := \langle \cdot, \cdot \rangle^{1/2}$ .

**ЧИСТО ЖАДНЫЙ АЛГОРИТМ (ЧЖА)** Пусть задан словарь  $\mathcal{D} \subset H$  и целевая функция  $f_0^{PGA} := f_0 := f \in H$ . Положим  $G_0^{PGA}(f, \mathcal{D}) := 0$ . Последовательно для каждого  $m \geq 0$  выбираем вектор  $g_{m+1} \in \mathcal{D}$  такой, что

$$|\langle f_m, g_{m+1} \rangle| = \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle|, \quad (2)$$

и определяем

$$G_{m+1}^{PGA}(f, \mathcal{D}) := G_m^{PGA}(f, \mathcal{D}) + \langle f_m, g_{m+1} \rangle g_{m+1},$$

$$f_{m+1}^{PGA} := f_{m+1} := f - G_{m+1}^{PGA}(f, \mathcal{D}) = f_m - \langle f_m, g_{m+1} \rangle g_{m+1}.$$

Таким образом, для  $f$  и  $m \geq 1$  построены  $m$ -членные приближения  $G_m^{PGA}(f, \mathcal{D})$ .

Корректность определения  $g_{m+1}$  по формуле (2) будет обсуждаться ниже, пока лишь заметим, что в случае конечномерного  $H$  и конечного словаря  $\mathcal{D}$  (случай приложений), супремум всегда достигается на каком-либо элементе словаря, и его вычисление требует конечного числа действий.

Отметим, что ЧЖА строит разложение элемента  $f$  в ряд по словарю  $\mathcal{D}$ , максимально уменьшая на каждом шаге алгоритма норму остатка:

$$\|f_{m+1}\| = \inf_{c \in \mathbb{R}, g \in \mathcal{D}} \|f_m - cg\|, \quad m \geq 0.$$

Если  $\mathcal{D}$  является ортонормированным базисом в  $H$ , то ЧЖА будет выбирать элементы словаря в порядке убывания коэффициентов Фурье функции  $f$  относительно этого базиса. Подобные приближения рассматривались в теории функций, начиная с работы С.Б.Стечкина [9].

Для переполненных систем специального вида ЧЖА впервые появился в статистике в работе Дж. Фридмана и В. Стузла [20] под именем Projection pursuit regression, а также в теории передачи сигнала в работе С. Маллата и Ж. Жанга [26] под именем Matching pursuit.

Необходимо также отметить, что результаты из ряда более ранних работ по теории функций, например, Е. Шмидта [27], а также Б.С. Стечкина и С.Б. Стечкина [8], могут быть проинтерпретированы как доказательства сходимости ЧЖА для словарей специального вида.

Жадные алгоритмы тесно связаны с орторекурсивными разложениями, которые в последние годы интенсивно исследовались Т.П. Лукашенко, В.В. Галатенко и др. ([4], [1]).

С современным состоянием теории жадных алгоритмов можно ознакомиться, например, в обзоре В.Н. Темлякова [29].

**Цель работы.** Получить оценки на скорость сходимости чисто жадного и ортогонального жадного алгоритмов для различных классов целевых функций. Изучить сходимость жадных алгоритмов для словарей с малой когерентностью. Разработать новые виды жадных алгоритмов,

позволяющие строить  $m$ -членные приближения с заданными свойствами. Исследовать аппроксимационные свойства “банаховых аналогов” чисто жадного алгоритма.

**Методы исследования.** В работе используются методы функционального анализа, теории приближения, геометрии банаховых пространств, теории функций действительного переменного, дискретной и вычислительной математики.

**Научная новизна и основные результаты.** Основные результаты диссертации являются новыми и состоят в следующем:

1. Показано, что чисто жадный алгоритм не является оптимальным по порядку в пространстве  $\mathcal{A}_1(\mathcal{D})$ . Получены оценки снизу на скорость сходимости чисто жадного алгоритма в пространствах  $\mathcal{A}_0(\mathcal{D})$  и  $\mathcal{A}_1(\mathcal{D})$ , достаточно близкие к наилучшим известным оценкам сверху (А.В. Сильбиченко). Тем самым получен ответ на вопрос Девора - Темлякова об оптимальности ЧЖА, и с точностью до 0.01 определена константа Колягина - Темлякова. (Теорема 1.12.)
2. Показано, что чисто жадный алгоритм обладает оптимальной по порядку скоростью сходимости в интерполяционных пространствах  $[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$  при  $0 < \theta < 1/3$ . (Теорема 1.15.)
3. Получено точное по порядку неравенство типа Лебега для ортогонального жадного алгоритма по словарям с малой когерентностью. Ранее эта задача исследовалась в работах Д. Донохо, М. Элада, В.Н. Темлякова, А. Гильберт, М. Мутукришнана, Дж. Штраусса, Дж. Троппа, П. Желтова. (Теорема 2.7.)
4. Предложен возвратный жадный алгоритм, который позволяет получать жадные разложения, обладающие оптимальной по порядку скоростью в интерполяционных пространствах  $[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$  при  $0 < \theta \leq 1$ , и, в частности, в  $\mathcal{A}_1(\mathcal{D})$ . (Теоремы 3.3, 3.4.)
5. Предложены положительный чисто жадный и положительный слабый жадный алгоритмы. Доказано, что если функция приближается элементами словаря с положительными коэффициентами, то жадные разложения, построенные с помощью этих алгоритмов, после приведения подобных будут иметь неотрицательные коэффициенты. Тем самым получен ответ на вопрос Б.С. Кашина о

конструктивном получении “положительных”  $m$ -членных приближений. (Теоремы 3.5, 3.6.)

6. Построен пример гладкого банахова пространства, словаря в нем и целевой функции, для которых  $X$ -жадный алгоритм расходится. (Теорема 4.1.)
7. Найдено новое геометрическое свойство банаховых пространств, являющееся достаточным для сходимости некоторых видов жадных алгоритмов в банаховых пространствах. Доказана сходимость  $R$ -жадного алгоритма в пространствах  $l^p$ ,  $p > 2$ . (Теоремы 4.2, 4.3.)
8. Исследована сходимость  $X$ -жадного алгоритма в пространстве  $L_p(0, 1)$  для конкретных систем: системы Хаара и системы функций, пропорциональных индикаторам двоичных интервалов. Получены неравенства типа Лебега, а также доказана сходимость  $X$ -жадного алгоритма по “системе индикаторов” на всем пространстве  $L_p(0, 1)$ ,  $1 < p < \infty$ . (Теоремы 4.4, 4.5, 4.6, 4.7.)

**Теоретическая и практическая ценность.** Работа носит теоретический характер. Полученные результаты и разработанные методы могут быть использованы в теории жадных алгоритмов, в теории приближения, вычислительной математике, теории передачи сигнала (в задаче сжатых измерений). Некоторые из разработанных в диссертации жадных алгоритмов могут оказаться полезными и в практических вычислительных задачах.

**Публикации.** Основные результаты диссертации опубликованы в работах [40]–[49] в изданиях, рекомендованных ВАК. В работах [33]–[38] опубликованы результаты автора о сходимости жадных алгоритмов, которые не были включены в диссертацию.

**Апробация работы.** Результаты диссертации докладывались в 1999–2010 годах на семинарах в МГУ: семинаре п/р П.Л. Ульянова и Б.С. Кашина, семинаре п/р Б.И. Голубова, М.С. Дьяченко, Б.С. Кашина и С.В. Конягина, семинаре п/р Б.С. Кашина и С.В. Конягина, семинаре п/р И.Г. Царькова, семинаре п/р Т.П. Лукашенко, семинаре п/р В.М. Тихомирова; в МИАНе на семинаре п/р Б.С. Кашина, на семинарах В РУДН: семинаре п/р А.В. Арутюнова, семинаре п/р А.Л. Скубачевского; на школах С.Б. Стечкина по теории приближения (Мнасс 2000,

2001, 2004, 2010гг., Алексин, 2007г.); на пленарных заседаниях конференции “Современные проблемы теории функций и приложения (Саратов, 2009г.), конференции “Конструктивная теория функция-2010” (Созополь, Болгария, 2010г.); на секционных заседаниях международных конференций “Теория приближения функций и операторов”, посвященной 80-летию С.Б. Стечкина (Екатеринбург, 2000г.), “Приближение и вероятность” (Бендлево, Польша, 2004г.), “Неортогональные разложения и жадные алгоритмы” (Вена, Австрия, 2005г.), “Кривые и поверхности” (Авиньон, Франция, 2006г.), “Симпозиум фон Неймана по разреженным представлениям и многомерной геометрии” (Сноуберд, США, 2007г.), “Современные проблемы математики, механики и их приложений”, посвященной 70-летию академика В.А. Садовниченко (Москва, 2009г.), “Современные проблемы анализа и преподавания математики”, посвященной 105-летию академика С.М. Никольского (Москва, 2010г.), “Теория приближения”, посвященной 90-летию С.Б. Стечкина (Москва, 2010г.), а также XXIV конференции молодых ученых механико-математического факультета МГУ им М.В. Ломоносова (Москва, 2002г.).

**Структура и объем работы.** Диссертация состоит из введения, четырех глав и списка литературы. Главы разбиты на параграфы. Некоторые параграфы разбиты на пункты. Нумерация утверждений (теоремы и леммы) двойная: номер главы и собственный номер. Нумерация формул тройная: номер главы, номер параграфа и собственный номер. Номера теорем во введении и автореферате совпадают с их номерами в основном тексте. Общий объем работы — 215 страниц. Список литературы содержит 106 наименований.

## Основное содержание работы.

Введение содержит краткую историю вопроса, описание целей и методов исследования и формулировки основных результатов диссертации.

В главе 1 изучается скорость сходимости чисто жадного и ортогонального жадного алгоритмов в гильбертовом пространстве. В параграфе 1.1 приводится определение чисто жадного и ортогонального жадного алгоритмов.

**ОРТОГОНАЛЬНЫЙ ЖАДНЫЙ АЛГОРИТМ (ОЖА)** Пусть задан словарь  $\mathcal{D}$  и целевая функция  $f_0^{OGA} := f_0 := f \in H$ . Положим  $G_0^{OGA}(f, \mathcal{D}) := 0$ . Последовательно для каждого  $t \geq 0$  выбираем век-



тор  $g_{m+1} \in \mathcal{D}$  такой, что

$$|\langle f_m, g_{m+1} \rangle| = \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle| \quad (3)$$

и определяем

$$G_{m+1}^{OGA}(f, \mathcal{D}) := \text{Proj}_{\text{span}(g_1, \dots, g_{m+1})} f,$$

$$f_{m+1}^{OGA} := f_{m+1} := f_0 - G_{m+1}^{OGA}(f, \mathcal{D}).$$

Также в параграфе 1.1 приводятся результаты Л. Джонса [24] и В. Дубинина [19] о сходимости ЧЖА и ОЖА: для любого гильбертова пространства  $H$ , словаря  $\mathcal{D}$  и целевой функции  $f \in H$  выполняются равенства

$$\lim_{m \rightarrow \infty} \|f - G_m^{PGA}(f, \mathcal{D})\| = 0, \quad \lim_{m \rightarrow \infty} \|f - G_m^{OGA}(f, \mathcal{D})\| = 0.$$

**Определение 1.1.** Будем называть словарь  $\mathcal{D}$  дискретным, если

$$\sup_{g_1, g_2 \in \mathcal{D}, g_1 \neq g_2} |\langle g_1, g_2 \rangle| < 1.$$

Легко видеть, что в сепарабельном пространстве дискретные словари не более чем счетны.

Формально супремум в формулах (2) и (3) может не достигаться. Однако оказалось, что если словарь  $\mathcal{D}$  дискретен, то для “большинства” функций супремум в формулах (2) и (3) достигается на каком-либо элементе словаря на всех шагах алгоритма.

**Теорема 1.1.** Пусть задан дискретный словарь  $\mathcal{D}$ . Тогда множество функций из  $H$ , для которых на каждом шаге ЧЖА супремум достигается:

$$\left\{ f \in H \mid \forall m \{g_{m+1} \in \mathcal{D} : |\langle f_m, g_{m+1} \rangle| = \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle|\} \neq \emptyset \right\}$$

является множеством второй категории<sup>1</sup> (использовалось обозначение  $f_m = f - G_m^{PGA}(f, \mathcal{D})$ ).

<sup>1</sup>под множествами второй категории в этой работе понимаются такие множества, дополнения которых являются множествами первой категории.

**Теорема 1.2.** Пусть задан дискретный словарь  $\mathcal{D}$ . Тогда множество функций из  $H$ , для которых на каждом шаге ОЖА супремум достигается:

$$\left\{ f \in H \mid \forall m \{g_{m+1} \in \mathcal{D} : |\langle f_m, g_{m+1} \rangle| = \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle|\} \neq \emptyset \right\}$$

является множеством второй категории (использовалось обозначение  $f_m = f - G_m^{OGA}(f, \mathcal{D})$ ).

В параграфе 1.3 приводятся результаты о скорости сходимости жадных алгоритмов. Для получения оценок сверху на скорость сходимости жадного алгоритма, необходимо потребовать, чтобы целевая функция  $f$  принадлежала некоторому, зависящему от словаря, классу  $\mathcal{A}(\mathcal{D})$ . Показано, что классы, определяемые скоростью убывания наилучшего  $m$ -членного приближения, оказываются слишком широкими.

**Теорема 1.5.** Пусть заданы строго убывающие положительные функции  $r, R \in C(\mathbb{R}_+)$ , для которых

$$\lim_{t \rightarrow \infty} r(t) = \lim_{t \rightarrow \infty} R(t) = 0.$$

Тогда найдется словарь  $\mathcal{D}$ , целевая функция  $f \in H$  и  $C > 0$ , для которых имеют место неравенства

$$\sigma_m(f, \mathcal{D}) \leq R(m), \quad m \geq 1,$$

$$\|f - G_m^{PGA}(f, \mathcal{D})\| \geq Cr(m), \quad m \geq 1,$$

$$\|f - G_m^{OGA}(f, \mathcal{D})\| \geq Cr(m), \quad m \geq 1.$$

Таким образом, для получения нетривиальных оценок сверху на скорость сходимости жадных алгоритмов необходимо накладывать более жесткие ограничения на класс  $\mathcal{A}(\mathcal{D})$ . Наиболее популярным в теории жадных алгоритмов является класс  $\mathcal{A}_1(\mathcal{D})$ , определенный ниже.

Для словаря  $\mathcal{D}$  и  $M > 0$  рассмотрим класс

$$\mathcal{A}_1^o(\mathcal{D}, M) := \left\{ f \in H : f = \sum_{\lambda \in \Lambda} c_\lambda g_\lambda, \quad g_\lambda \in \mathcal{D}, \quad \#\Lambda < \infty \text{ и } \sum_{\lambda \in \Lambda} |c_\lambda| \leq M \right\}$$

и класс  $\mathcal{A}_1(\mathcal{D}, M)$ , являющийся замыканием (в  $H$ ) класса  $\mathcal{A}_1^o(\mathcal{D}, M)$ . Далее определим

$$\mathcal{A}_1(\mathcal{D}) := \bigcup_{M \geq 0} \mathcal{A}_1(\mathcal{D}, M).$$

Для  $f \in \mathcal{A}_1(\mathcal{D})$ , введем норму

$$|f|_{\mathcal{A}_1(\mathcal{D})} = \inf\{M \geq 0 \mid f \in \mathcal{A}_1(\mathcal{D}, M)\}.$$

Отметим, что шар  $\mathcal{A}_1(\mathcal{D}, 1)$  является замыканием выпуклой оболочки  $\mathcal{D} \cup (-\mathcal{D})$ .

Из результатов С.Б. Стечкина [9] и А. Бэррона [10] следует, что последовательность величин наилучшего  $m$ -членного приближения  $\sigma_m(f, \mathcal{D})$  для  $f \in \mathcal{A}_1(\mathcal{D})$  убывает не медленнее, чем  $Cm^{-1/2}$ , и показатель  $-1/2$  не может быть уменьшен даже для ортонормированного словаря.

Приводятся оценки сверху на скорость сходимости ЧЖА

$$\|f - G_m^{PGA}(f, \mathcal{D})\| \leq C_\gamma |f|_{\mathcal{A}_1(\mathcal{D})} m^{-\gamma}. \quad (4)$$

Р. ДеВор и В.Н. Темляков [14] доказали справедливость (4) для  $\gamma = 1/6$ , С.В. Конягин и В.Н. Темляков [25] — для  $\gamma = 11/62$ , А.В. Сильиченко [7] — для  $\gamma = 0.182$ .

Для построения нижних оценок необходимо построить словарь  $\mathcal{D}$ , функцию  $f \in \mathcal{A}_1(\mathcal{D})$ ,  $f \neq 0$  и для нее оценить снизу скорость убывания нормы  $\|f - G_m^{PGA}(f, \mathcal{D})\|$ :

$$\|f - G_m^{PGA}(f, \mathcal{D})\| \geq C_\gamma |f|_{\mathcal{A}_1(\mathcal{D})} m^{-\gamma}, \quad C_\gamma > 0. \quad (5)$$

Отметим, что во всех нижних оценках, приведенных в работе, функция  $f$  будет являться конечной линейной комбинацией элементов словаря, т.е. принадлежать пространству разреженных (sparse) векторов  $\mathcal{A}_0(\mathcal{D})$ :

$$\Sigma_m(\mathcal{D}) = \left\{ \sum_{\lambda \in \Lambda} c_\lambda g_\lambda, \quad c_\lambda \in \mathbb{R}, \quad g_\lambda \in \mathcal{D}, \quad \#\Lambda < m \right\}. \quad (6)$$

$$\mathcal{A}_0(\mathcal{D}) := \bigcup_{m \geq 0} \Sigma_m(\mathcal{D}),$$

$$|f|_0 = |f|_{\mathcal{A}_0(\mathcal{D})} := \min\{m \geq 0 : f \in \Sigma_m(\mathcal{D})\}, \quad f \in \mathcal{A}_0(\mathcal{D}).$$

Класс  $\mathcal{A}_0(\mathcal{D}) \subset \mathcal{A}_1(\mathcal{D})$  является очень узким и из приведенных ниже оценок будет следовать, что за счет “разумного” сужения  $\mathcal{A}_1(\mathcal{D})$  существенно увеличить скорость сходимости ЧЖА и ОЖА нельзя.

Р. ДеВор и В.Н. Темляков [14] установили (5) для  $\gamma = 1/2$ . Автору удалось проверить (5) для  $\gamma = 1/2 - \delta$ ,  $\delta > 0$ , В.Н. Темляков уменьшил  $\gamma$  до  $1/3$ , в совместной работе автора и В.Н. Темлякова [36]  $\gamma$  было уменьшено до 0.27.

В параграфе 1.4 получена новая оценка снизу на скорость сходимости ЧЖА, которая оказывается весьма близкой к верхней оценке А.В. Сильниченко.

**Теорема 1.12.** *Существует такой словарь  $\mathcal{D}$ , элемент  $f \in \mathcal{A}_0(\mathcal{D})$  и  $C > 0$ , что для произвольного  $m \geq 1$  выполняется неравенство*

$$\|f - G_m^{PGA}(f, \mathcal{D})\| = \|f_m^{PGA}\| > Cm^{-0.1898} \|f\|_{\mathcal{A}_0(\mathcal{D})}.$$

В параграфе 1.5 получена оценка снизу на скорость сходимости ортогонального жадного алгоритма.

**Теорема 1.13.** *Существует такой словарь  $\mathcal{D}$ , элемент  $f \in \mathcal{A}_0(\mathcal{D})$  и  $C > 0$ , что для произвольного  $m \geq 1$  выполняется неравенство*

$$\|f - G_m^{OGA}(f, \mathcal{D})\| > Cm^{-1/2}.$$

Этот результат показывает неулучшаемость оценки сверху Р. ДеВора и В.Н. Темлякова [14] на скорость сходимости ОЖА.

Как отмечалось выше, в случае произвольного словаря  $\mathcal{D}$  сужение класса  $\mathcal{A}_1(\mathcal{D})$  не ведет к увеличению скорости сходимости жадных алгоритмов, в тоже время как на более широких классах, скорость сходимости жадных алгоритмов может быть близка к оптимальной. Соответствующие результаты формулируются в параграфе 1.6 и доказываются в параграфе 1.7.

В качестве расширений  $\mathcal{A}_1(\mathcal{D})$  будут рассматриваться интерполяционные пространства  $[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$ .

Пусть  $X$  и  $Y$  действительные нормированные пространства. Напомним (см. [13]) определение интерполяционных пространств  $[X, Y]_{\theta, \infty}$ ,  $0 < \theta < 1$ . По определению  $f \in X$  принадлежит пространству  $[X, Y]_{\theta, \infty}$  тогда и только тогда, если существует такое  $C > 0$ , что для всех  $t > 0$  выполняется

$$K(f, t) \leq Ct^\theta, \tag{7}$$

где

$$K(f, t) = K(f, t, X, Y) := \inf_{h \in Y} \{\|f - h\|_X + t\|h\|_Y\}.$$

В качестве нормы  $\|f\|_{[X,Y]_{\theta,\infty}}$ , берется инфимум по числам  $C$ , для которого выполняется неравенство (7).

А. Бэррон, А. Коэн, В. Дамен, и Р. ДеВор [12] доказали оптимальность ОЖА в интерполяционных пространствах  $[H, \mathcal{A}_1(\mathcal{D})]_{\theta,\infty}$ , для всех  $\theta$ ,  $0 < \theta \leq 1$ . В случае малых  $\theta$  чисто жадный алгоритм также является почти оптимальным:

**Теорема 1.15.** *Для произвольных  $\theta$ ,  $0 < \theta \leq 1/3$  и  $\epsilon$ ,  $0 < \epsilon < \theta/2$  существует такое  $C > 0$  что для всех  $f \in [H, \mathcal{A}_1(\mathcal{D})]_{\theta,\infty}$  и  $m \geq 1$  справедлива оценка*

$$\|f - G_m^{PGA}(f, \mathcal{D})\| \leq C \|f\|_{[H, \mathcal{A}_1(\mathcal{D})]_{\theta,\infty}} m^{-\theta/2 + \epsilon}.$$

В параграфе 1.8 доказываются оценки снизу на скорость убывания наилучшего  $m$ -членного приближения для интерполяционных классов.

Из теорем 1.12 и 1.13 вытекает, что если не накладывать на словарь  $\mathcal{D}$  никаких дополнительных ограничений, то жадные алгоритмы не могут гарантировать скорость сходимости, лучшую чем  $Cm^{-1/2}$  даже для разреженных целевых функций (функций из  $\mathcal{A}_0(\mathcal{D})$ ). Тем не менее оказалось, что при наложении на словарь некоторых ограничений, связанных с когерентностью, скорость сходимости жадных алгоритмов может быть значительно выше.

В главе 2 исследуется скорость сходимости жадных алгоритмов по словарям с малой когерентностью. В начале параграфа 2.1 приводится определение когерентности, указывается пример переполненного словаря с малой когерентностью, построенного на основе Бернуллиевской матрицы, дается краткая историческая справка и приводятся формулировки основных результатов.

**Определение 2.1.** Когерентностью словаря  $\mathcal{D}$  будем называть величину

$$\mu := \sup_{\phi, \psi \in \mathcal{D}, \phi \neq \psi} |\langle \phi, \psi \rangle|.$$

Словари с когерентностью  $\mu$  будем называть  $\mu$ -когерентными.

Сходимость ортогонального жадного алгоритма по словарям с малой когерентностью изучалась в работах Р. Грибонваля, М. Нильсена, Д. Донохо, М. Элада, В.Н. Темлякова, А. Гильберт, М. Мутукришнапа, Дж. Штраусса, Дж. Троппа и др. Интерес к этому направлению в

значительной степени обусловлен тем, что получаемые здесь результаты оказались востребованы в задаче “сжатых измерений” (compressed sensing).

Следующая теорема является базовым результатом о сходимости ортогонального жадного алгоритма для разреженных целевых функций по словарю с малой когерентностью.

**Теорема 2.1** ([23], [22], [32]) Пусть задан словарь  $\mathcal{D}$  с когерентностью  $\mu$  и функция  $f \in A_0(\mathcal{D})$ . Если

$$|f|_0 < \frac{1}{2}(\mu^{-1} + 1), \quad (8)$$

то для любого  $m \geq |f|_0$  имеет место равенство

$$f = G_m^{OGA}(f, \mathcal{D}).$$

Как было показано В.Н. Темляковым и П. Желтовым [31], константа  $1/2$  в неравенстве (8) не может быть увеличена.

Устойчивость точной аппроксимации разреженных целевых функций интенсивно изучалась. Следуя В.Н. Темлякову, будем называть *неравенствами типа Лебега* оценки, связывающие ошибку аппроксимации жадного алгоритма и наилучшее  $m$ -членное приближение

$$\|f - G_{A(m)}^{OGA}(f, \mathcal{D})\| \leq B(m)\sigma_m(f, \mathcal{D}), \quad m \leq C(\mu). \quad (9)$$

А. Гильберт, М. Мутукришнан и Дж. Штраусс [22] установили неравенство (9) для  $A(m) = m$ ,  $B(m) = 8m^{1/2}$ ,  $C(\mu) = 2^{-3.5}\mu^{-1} - 1$ . Константы в  $A(m)$  и  $B(m)$  были несколько улучшены Дж. Троппом [32] (см. также статью Д. Л. Донохо, М. Элада и В.Н. Темлякова [17].) Позднее Д. Л. Донохо, М. Элад и В.Н. Темляков [18] значительно улучшили  $B(m)$  и доказали неравенство (9) с  $A(m) = \lfloor m \log m \rfloor$ ,  $B(m) = 24$ ,  $C(\mu) = \frac{1}{26}\mu^{-2/3}$ . В своей недавней работе [31] В.Н. Темляков и П. Желтов приблизили значение  $C(\mu)$  к оптимальному, доказав оценку (9) с  $A(m) = m \lfloor 2^{\sqrt{\log m}} \rfloor$ ,  $B(m) = 3$  и  $C(\mu)$ , которое должно обеспечить неравенство  $m 2^{\sqrt{2 \log m}} \leq \frac{1}{26}\mu^{-1}$  для  $m \leq C(\mu)$ .

В параграфе 2.1 получено точное по порядку неравенство типа Лебега (для ортогонального жадного алгоритма).

**Теорема 2.7** Для произвольного  $\mu$ -когерентного словаря  $\mathcal{D}$  и функции  $f \in H$  справедлива оценка

$$\|f - G_{2m}^{OGA}(f, \mathcal{D})\| \leq 3\sigma_m(f)$$

для всех

$$1 \leq m \leq \frac{1}{20\mu}.$$

Этот результат доказывается в пунктах 2.1.1 – 2.1.4.

В параграфе 2.2 исследуются словари с ограниченной совокупной когерентностью. Следуя Дж. Троппу [32], введем определение.

**Определение 2.2.** Совокупной когерентностью счетного словаря  $\mathcal{D}$  будем называть величину

$$\mu_1(\mathcal{D}) := \sup_{g \in \mathcal{D}} \sum_{\tilde{g} \in \mathcal{D}, \tilde{g} \neq g} |\langle \tilde{g}, g \rangle|.$$

Будем говорить, что словарь  $\mathcal{D}$  имеет ограниченную совокупную когерентность, если для него  $\mu_1(\mathcal{D}) < \infty$ .

Хорошо известно, что если для словаря  $\mathcal{D}$  имеет место оценка  $\mu_1(\mathcal{D}) < 1/2$ , то для него выполняется так называемое trapping-свойство. Следующий результат в том или ином виде присутствует во всех работах по скорости сходимости жадных алгоритмов, в которых накладываются ограничения на когерентность словаря.

**Теорема 2.8 (trapping-свойство, [23], [17], [22], [32])** Пусть заданы словарь  $\mathcal{D}$  с  $\mu_1(\mathcal{D}) < 1/2$ , его конечное подмножество  $\mathcal{D}_0$  и  $f \in \text{span}(\mathcal{D}_0)$ ,  $f \neq 0$ . Тогда

$$\max_{g \in \mathcal{D}_0} |\langle f, g \rangle| > \sup_{g \in \mathcal{D} \setminus \mathcal{D}_0} |\langle f, g \rangle|.$$

Теорема 2.8 гарантирует, что если  $f \in \text{span}(\mathcal{D}_0)$ , где  $\mathcal{D}_0$  конечное подмножество  $\mathcal{D}$ , то на каждом шаге жадного алгоритма очередной элемент словаря будет выбираться из  $\mathcal{D}_0$ . Тем самым, в случае ортогонального жадного алгоритма за конечное число шагов ( $\#\mathcal{D}_0$ ) будет получено точное решение, а в случае чисто жадного алгоритма будет иметь место экспоненциальная скорость сходимости (как в конечномерных пространствах).

При малых значениях  $\mu_1(\mathcal{D})$  чисто жадный алгоритм обладает оптимальной скоростью и в пространстве  $\mathcal{A}_1(\mathcal{D})$

**Теорема 2.9.** Пусть заданы словарь  $\mathcal{D}$  с  $\mu_1(\mathcal{D}) < 1/3$  и  $f \in \mathcal{A}_1(\mathcal{D})$ . Тогда

$$\|f - G_m^{PGA}(f, \mathcal{D})\| \leq |f|_1 m^{-1/2}, \quad m \geq 1.$$

Теорема 2.9 была обобщена на более широкие классы.  
 Определим классы  $\mathcal{A}^p(\mathcal{D})$ ,  $1 \leq p < \infty$ . Положим для  $M \geq 0$

$$\mathcal{A}^p(\mathcal{D}, M) := \overline{\left\{ \sum_{\lambda \in \Lambda} c_\lambda g^\lambda : \sum_{\lambda \in \Lambda} |c_\lambda|^p \leq M^p, c_\lambda \in \mathbb{R}, g^\lambda \in \mathcal{D}, \#\Lambda < \infty \right\}},$$

где замыкание берется в норме пространства  $H$ . Пусть

$$\mathcal{A}^p(\mathcal{D}) := \bigcup_{M \geq 0} \mathcal{A}^p(\mathcal{D}, M),$$

$$\|f\|_p := |f|_{\mathcal{A}^p(\mathcal{D})} := \inf\{M \geq 0 : f \in \mathcal{A}^p(\mathcal{D}, M)\}, f \in \mathcal{A}^p(\mathcal{D}).$$

**Теорема 2.10.** Пусть заданы словарь  $\mathcal{D}$  с  $\mu_1(\mathcal{D}) < 1/3$ ,  $p$ ,  $1 \leq p < 2$  и  $f \in \mathcal{A}^p(\mathcal{D})$ . Тогда найдутся  $C_1 = C_1(p) > 0$  и  $C_2 = C_2(\mu_1(\mathcal{D})) > 0$  такие, что для любого  $m \geq 1$

$$\|f - G_m^{PGA}(f, \mathcal{D})\| \leq C_1 C_2 \|f\|_p m^{-1/p+1/2}.$$

В главе 3 исследуется вопрос о получении для целевой функции  $f$  разложений по элементам словаря  $\mathcal{D}$  вида

$$f \sim \sum_{k=1}^{\infty} c_k(f) g_k(f), c_k(f) \in \mathbb{R}, g_k(f) \in \mathcal{D},$$

частичные суммы которых

$$\sum_{k=1}^m c_k(f) g_k(f)$$

достаточно хорошо приближают  $f$  (при этом различность  $g_k(f)$  не требуется). Легко видеть, что чисто жадному алгоритму соответствует так называемое ЧЖА-разложение

$$f \sim \sum_{k=1}^{\infty} \langle f_{k-1}^{PGA}, g_k^{PGA} \rangle g_k^{PGA},$$

в то время как ортогональному жадному алгоритму, в силу того, что на каждом шаге идет полный пересчет коэффициентов, никакое разложение не соответствует.



В главе 3 будут построены разложения, частичные суммы которых на классах  $\mathcal{A}_1(\mathcal{D})$  и  $[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$  обеспечивают скорость приближения близкую к оптимальной.

Первые разложения, превосходящие по скорости ЧЖА-разложения были предложены В.Н. Темляковым в работе [30] (см. также [29]). Им был рассмотрен слабый жадный алгоритм с параметром  $b$  и получены оценки на его скорость сходимости

**СЛАБЫЙ ЖАДНЫЙ АЛГОРИТМ С ПАРАМЕТРОМ  $b$  (СЖАВ)** Пусть заданы последовательность  $\{t_m\}_{m=1}^{\infty} \subset [0, 1]$ , параметр  $b$ ,  $0 < b \leq 1$ , словарь  $\mathcal{D}$  и целевая функция  $f_0^{WGAb} := f_0 := f \in H$ . Положим  $G_0^{WGAb}(f, \mathcal{D}) := 0$ . Последовательно для каждого  $m \geq 0$  выбираем вектор  $g_{m+1}^{WGAb} := g_{m+1} \in \mathcal{D}$  такой, что

$$|\langle f_m, g_{m+1} \rangle| \geq t_{m+1} \sup_{g \in \mathcal{D}} |\langle f_m, g \rangle|$$

и определяем

$$G_{m+1}^{WGAb}(f, \mathcal{D}) := G_m^{WGAb}(f, \mathcal{D}) + b \langle f_m, g_{m+1} \rangle g_{m+1},$$

$$f_{m+1}^{WGAb} := f_{m+1} := f - G_{m+1}^{WGAb}(f, \mathcal{D}) = f_m - b \langle f_m, g_{m+1} \rangle g_{m+1}.$$

В.Н. Темляков [30] доказал, что для произвольного элемента  $f \in \mathcal{A}_1(\mathcal{D})$  и  $m \geq 1$  справедлива оценка

$$\|f - G_m^{WGAb}(f, \mathcal{D})\| \leq \|f\|_{\mathcal{A}_1(\mathcal{D})} (1 + b(2-b)) \sum_{k=1}^m t_k^2 \frac{-(2-b)t_m}{2(2+(2-b)t_m)}. \quad (10)$$

Легко видеть, что если для достаточно малого  $\epsilon > 0$  положить

$$b = \epsilon, \quad t_m = 1 - \epsilon, \quad m \geq 1,$$

то порядок скорости сходимости СЖАВ окажется близким к  $-1/4$ , т.е. существенно лучшим по сравнению со скоростью сходимости ЧЖА.

Перейдем к определению *возвратного жадного алгоритма*, который, с одной стороны, как чисто жадный алгоритм дает разложение целевой функции в ряд по элементам словаря, а с другой, как ортогональный жадный алгоритмы, дает правильную в полиномиальной шкале скорость сходимости в классах  $\mathcal{A}_1(\mathcal{D})$  и  $[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$ ,  $0 < \theta \leq 1$ .

**ВОЗВРАТНЫЙ ЖАДНЫЙ АЛГОРИТМ (ВЖА)** Пусть заданы вектор  $f \in H$ , словарь  $\mathcal{D}$  и число  $0 < t \leq 1$ . Положим  $f_0^{RecGA} := f_0 = f$ ,

$G_0^{RecGA}(f, \mathcal{D}, t) := G_0(f, \mathcal{D}, t) := 0$ . Предположим, что для  $n \geq 0$  уже определены  $f_1, \dots, f_n \in H$ ,  $g_1, \dots, g_n \in \mathcal{D}$  и  $c_1, \dots, c_n \in \mathbb{R}$ . Причем

$$f_k = f - \sum_{i=1}^k c_i g_i, \quad \forall k, 1 \leq k \leq n. \quad (11)$$

Обозначим через

$$D_n := D_n(f_0, \mathcal{D}, t) := \{g_k\}_{k=1}^n; \quad (12)$$

$$v_n(g) := v_n(g, f_0, \mathcal{D}, t) := \sum_{k \leq n, g_k = g} c_k; \quad (13)$$

$$b_n := b_n(f_0, \mathcal{D}, t) := \sum_{g \in D_n} |v_n(g)|;$$

$$a_n = \langle f_n, f_n \rangle.$$

Если  $n \geq 1$  и существует хотя бы один вектор  $g \in D_n$  такой, что выполняются неравенства

$$\min(|v_n(g)|, |\langle f_n, g \rangle|) \geq \frac{a_n t^2}{9b_n},$$

и

$$\text{sign}(v_n(g)) = -\text{sign}(\langle f_n, g \rangle),$$

то положим

$$g_{n+1}^{RecGA} := g_{n+1} := g, \quad c_{n+1} = \text{sign}(\langle f_n, g \rangle) \min(|v_n(g)|, |\langle f_n, g \rangle|).$$

Будем называть такие шаги возвратными. В противном случае, выберем произвольный  $g_{n+1}^{RecGA} = g_{n+1} \in \mathcal{D}$ , удовлетворяющий неравенству

$$|\langle f_n, g_{n+1} \rangle| \geq t \sup_{g \in \mathcal{D}} |\langle f_n, g \rangle|$$

и положим

$$c_{n+1} = \langle f_n, g_{n+1} \rangle$$

Будем называть такие шаги жадными. Определим аппроксимант и остаток:

$$G_{n+1}^{RecGA}(f, \mathcal{D}, t) := G_{n+1}(f, \mathcal{D}, t) := \sum_{k=1}^{n+1} c_k g_k = G_n(f, \mathcal{D}) + c_{n+1} g_{n+1}.$$

$$f_{n+1}^{RecGA} := f_{n+1} := f_n - c_{n+1}g_{n+1} = f - G_{n+1}(f, \mathcal{D}, t),$$

что гарантирует выполнение (11) на следующем шаге алгоритма.

Ряд

$$\sum_{n=1}^{\infty} c_n g_n^{RecGA}$$

будем называть возвратным жадным разложением элемента  $f$  по словарю  $\mathcal{D}$ .

В параграфе 3.2 доказывается следующая теорема о сходимости возвратных разложений.

**Теорема 3.2** Для произвольного словаря  $\mathcal{D}$ , целевой функции  $f \in H$  и  $t$ ,  $0 < t \leq 1$ , возвратный жадный алгоритм сходится:

$$\lim_{n \rightarrow \infty} \|f - G_n^{RecGA}(f, \mathcal{D}, t)\| = 0.$$

В параграфе 3.3 установлены следующие две теоремы о скорости сходимости ВЖА.

**Теорема 3.3.** Для произвольного  $t$ ,  $0 < t \leq 1$ , существует такая константа  $K = K(t) > 0$ , что для произвольного словаря  $\mathcal{D}$ , целевой функции  $f \in \mathcal{A}_1(\mathcal{D})$  и всех  $n \geq 1$  имеет место неравенство

$$\|f - G_n^{RecGA}(f, \mathcal{D}, t)\| \leq K |f|_{\mathcal{A}_1(\mathcal{D})} n^{-1/2} \ln n. \quad (14)$$

**Теорема 3.4.** Для произвольных чисел  $t \in (0, 1]$  и  $\theta \in (0, 1]$ , существует такая константа  $K = K(t, \theta)$ , что для любого словаря  $\mathcal{D}$ , целевой функции  $f \in [H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}$  и всех  $n \geq 1$  имеет место неравенство

$$\|f - G_n^{RecGA}(f, \mathcal{D}, t)\| \leq K |f|_{[H, \mathcal{A}_1(\mathcal{D})]_{\theta, \infty}} n^{-\theta/2} \ln n.$$

**Замечание 3.2.** В отличие от оценки (10) показатель степени в оценках скорости сходимости в теоремах 3.3 и 3.4 не зависит от параметра  $t$ .

Во многих приложениях, в частности задачах, относящихся к финансовой математике, требуется, чтобы коэффициенты  $c_k$  в  $n$ -членном приближении

$$f \sim \sum_{k=1}^n c_k g_k, \quad g_k \in \mathcal{D}, \quad (15)$$

были неотрицательными. В параграфе 3.4 обсуждается получение  $n$ -членных приближений с неотрицательными коэффициентами с помощью жадных алгоритмов.

Ясно, что при наложении дополнительного условия (неотрицательность коэффициентов) на  $n$ -членное приближение, мы должны также наложить дополнительное условие на словарь  $\mathcal{D}$ , чтобы построение соответствующего приближения было возможным. Множество  $\mathcal{D} \subset H$  будем называть положительным словарем, если

$$g \in \mathcal{D} \Rightarrow \|g\| = 1, \left\{ \sum_{\lambda \in \Lambda} c_\lambda g_\lambda \mid c_\lambda > 0, g_\lambda \in \mathcal{D}, \#\Lambda < \infty \right\} = H. \quad (16)$$

А. Вэрроном [11] был предложен способ построения неотрицательных  $n$ -членных приближений с помощью релаксационного жадного алгоритма. Из результатов работы [12] вытекает, что релаксационный жадный алгоритм сходится для произвольного положительного словаря  $\mathcal{D}$  и целевой функции  $f_0$ , этот алгоритм также обладает хорошей скоростью сходимости, оптимальной для интерполяционных классов. С другой стороны, релаксационный жадный алгоритм дает только  $n$ -членные приближения (15), но не дает разложения

$$f \sim \sum_{k=1}^{\infty} c_k g_k, \quad g_k \in \mathcal{D}, \quad (17)$$

кроме того, последовательность норм остатков не обязана монотонно убывать. В параграфе 3.4 мы рассматриваем “положительные аналоги” чисто жадного и слабого жадного алгоритмов, которые обеспечивают построение разложений (17), т.е. обладают свойством оперативности (online свойством), и частичные суммы которых (15) монотонно (в смысле нормы остатка) стремятся к целевой функции. (При этом с точки зрения скорости сходимости они могут уступать релаксационному жадному алгоритму).

При определении “положительных” версий ЧЖА и СЖА (ПЧЖА и ПСЖА) используются обозначения (12) и (13). Отметим, что ПЧЖА и ПСЖА не гарантируют неотрицательность всех коэффициентов  $c_k$  в разложении (17), но любая частичная сумма ряда может быть рассмотрена как  $n$ -членное приближение с неотрицательными коэффициентами.

ми.

$$\sum_{k=1}^n c_k g_k = \sum_{g \in \mathcal{D}_n} v_n(g)g; \quad v_n(g) \geq 0, \quad g \in \mathcal{D}_n.$$

**ПОЛОЖИТЕЛЬНЫЙ ЧИСТО ЖАДНЫЙ АЛГОРИТМ (ПЧЖА)** Пусть заданы целевая функция  $f_0 := f \in H$  и положительный словарь  $\mathcal{D}$ . Положим  $G_0^{PGA+}(f, \mathcal{D}) := 0$ . Последовательно для каждого  $m \geq 0$  выбираем вектор  $g_{m+1} \in \mathcal{D}$  и  $c_{m+1} \geq -v_m(g_{m+1})$  так, чтобы

$$\|f_m - c_{m+1}g_{m+1}\| = \inf_{g \in \mathcal{D}, c \geq -v_m(g)} \|f_m - cg\|,$$

и определяем

$$G_{m+1}^{PGA+}(f, \mathcal{D}) := G_m^{PGA+}(f, \mathcal{D}) + c_{m+1}g_{m+1}$$

и

$$f_{m+1} = f - G_{m+1}^{PGA+}(f, \mathcal{D}) = f_m - c_{m+1}g_{m+1}.$$

Справедлива следующая теорема о сходимости положительного чисто жадного алгоритма.

**Теорема 3.5.** Для произвольного положительного словаря  $\mathcal{D}$  и целевой функции  $f \in H$  выполняется

$$\lim_{m \rightarrow \infty} \|f - G_m^{PGA+}(f, \mathcal{D})\| = 0.$$

Также в параграфе 3.4 получено достаточное условие на коэффициенты слабости для сходимости положительного слабого жадного алгоритма.

**Глава 4** посвящена изучению сходимости и скорости сходимости жадных алгоритмов в банаховых пространствах. Это направление развивалось с конца 1990-х – начала 2000-х годов в работах В.Н. Темлякова, С. Дилворфа, Д. Куцаровой, С.В. Колягина, П. Войташческа, Н. Калтона, С.Л. Гогяна и др. для многих результатов о сходимости жадных алгоритмов в гильбертовом пространстве нашелся “банаховый” аналог.

В параграфе 4.1 обсуждаются различные обобщения чисто жадного алгоритма на случай произвольного банахова пространства.

Пусть  $X = (X, \|\cdot\|)$  - действительное банахово пространство, а  $D \subset X$  словарь в нем. Для  $f \in X$  обозначим через  $F_f$  опорный функционал к  $f$ :

$$F_f \in X^*, \|F_f\| = 1, F_f(f) = \|f\|. \quad (18)$$

Для  $f, g \in X$ ,  $\|g\| = 1$  определим

$$\mu(f, g) \in \mathbb{R} : \|f - \mu(f, g)g\| = \inf_{\mu} \|f - \mu g\|, \quad (19)$$

$$r_q(f, g) := \|f\|^q - \|f - \mu(f, g)g\|^q, \quad q > 0,$$

$$r(f, g) := r_1(f, g) = \|f\| - \|f - \mu(f, g)g\|.$$

Отметим, что согласно теореме Хаана-Банаха функционалы, удовлетворяющие (18), существуют и в качестве  $F_f$  можно выбрать любой их них. В случае, если пространство  $X$  является гладким (по Гато), то для любого  $f \in X$  выбор  $F_f$  становится однозначным. Аналогично, в качестве  $\mu(f, g)$  берется любое число, удовлетворяющее (19), которое в строго выпуклых пространствах определяется однозначно.

**$X$ -ЖАДНЫЙ АЛГОРИТМ ( $X$ -ЖА).** Пусть задан словарь  $\mathcal{D}$  и целевая функция  $f_0 := f \in H$ . Положим  $G_0^{XGA}(f, \mathcal{D}) := 0$ . Последовательно для каждого  $m \geq 0$  выбираем вектор  $g_{m+1} \in \mathcal{D}$  такой, что

$$r(f_m, g_{m+1}) = \sup_{g \in \mathcal{D}} r(f_m, g)$$

и определяем

$$G_{m+1}^{XGA}(f, \mathcal{D}) := G_m^{XGA}(f, \mathcal{D}) + \mu(f_m, g_{m+1})g_{m+1},$$

$$f_{m+1} := f - G_{m+1}^{XGA}(f, \mathcal{D}) = f_m - \mu(f_m, g_{m+1})g_{m+1}.$$

В. Дубинин [19] отметил, что гладкость пространства является необходимым условием для сходимости  $X$ -жадного алгоритма. В параграфе 4.2 доказывается, что гладкость пространства не является достаточным условием сходимости.

**Теорема 4.1.** Существует гладкое (по Гато) банахово пространство  $X$ , словарь  $\mathcal{D} \subset X$  и  $f \in X$ , для которых

$$\lim_{m \rightarrow \infty} \|f_m\| = \lim_{m \rightarrow \infty} \|f - G_m^{XGA}(f)\| \neq 0.$$

Рассмотрим другие обобщения ЧЖА.

**ДВОЙСТВЕННЫЙ ЖАДНЫЙ АЛГОРИТМ С ПАРАМЕТРОМ  $t$  ( $D$ -ЖА).** Пусть задан словарь  $\mathcal{D}$ , целевая функция  $f_0 := f \in X$  и фиксированное число  $t$ ,  $0 < t < 1$ . Положим  $G_0^{DGA}(f, \mathcal{D}, t) := 0$ . Последовательно для каждого  $m \geq 0$  выбираем вектор  $g_{m+1} \in \mathcal{D}$  такой, что

$$|F_{f_m}(g_{m+1})| \geq t \sup_{g \in \mathcal{D}} |F_{f_m}(g)|,$$

и определяем

$$G_{m+1}^{DGA}(f, \mathcal{D}, t) := G_m^{DGA}(f, \mathcal{D}, t) + \mu(f_m, g_{m+1})g_{m+1},$$

$$f_{m+1} := f - G_{m+1}^{DGA}(f, \mathcal{D}) = f_m - \mu(f_m, g_{m+1})g_{m+1}.$$

**$R$ -ЖАДНЫЙ АЛГОРИТМ С ПАРАМЕТРАМИ  $t$  и  $q$  ( $R$ -ЖА).** Пусть задан словарь  $\mathcal{D}$ , целевая функция  $f_0 := f \in X$  и фиксированные числа  $t$ ,  $0 < t \leq 1$  и  $q$ ,  $q > 0$ . Положим  $G_0^{RGA}(f, \mathcal{D}, t, q) := 0$ . Для каждого  $m \geq 0$  последовательно выбираем вектор  $g_{m+1} \in \mathcal{D}$  такой, что

$$\frac{r_q(f_m, g_{m+1})}{|\mu(f_m, g_{m+1})|} \geq t \sup_{g \in \mathcal{D}} \frac{r_q(f_m, g)}{|\mu(f_m, g)|}$$

и определяем

$$G_{m+1}^{RGA}(f, \mathcal{D}, t, q) := G_m^{RGA}(f, \mathcal{D}, t, q) + \mu(f_m, g_{m+1})g_{m+1},$$

$$f_{m+1} := f - G_{m+1}^{RGA}(f, \mathcal{D}, t, q) = f_m - \mu(f_m, g_{m+1})g_{m+1}.$$

Отметим, что, если  $X$  является гильбертовым пространством, то  $X$ -ЖА,  $D$ -ЖА с  $t = 1$  и  $R$ -ЖА с  $t = 1$ ,  $q = 2$  в точности совпадают с ЧЖА.

Выбор параметра  $q$  в формулировке  $R$ -жадного алгоритма может оказывать влияние на скорость сходимости, но сам факт сходимости от него не зависит:

**Лемма 4.1.** Пусть заданы банахово пространство  $X$  и словарь  $\mathcal{D} \subset X$ . Если существует  $q_0 > 0$ , что для всех  $0 < t \leq 1$ , всех  $f \in X$  и произвольной реализации  $R$ -жадного алгоритма выполняется

$$\lim_{m \rightarrow \infty} \|f - G_m^{RGA}(f, \mathcal{D}, t, q_0)\| = 0,$$

то для всех  $q > 0$ ,  $0 < t \leq 1$ ,  $f \in X$  и произвольной реализации  $R$ -жадного алгоритма имеет место равенство

$$\lim_{m \rightarrow \infty} \|f - G_m^{RGA}(f, \mathcal{D}, t, q)\| = 0.$$

Оказалось, что для сходимости жадных алгоритмов в равномерно выпуклых и равномерно гладких пространствах (где, в частности,  $\mu(f, g)$  и  $F_f(g)$  определены однозначно) важную роль играет следующее дополнительное геометрическое свойство:

$$\exists \Theta > 1 : \forall f, g \in X \|g\| = 1, \Theta r(f, g) \geq |\mu(f, g)F_f(g)|. \quad (20)$$

В параграфе 4.3 доказывается общая теорема о сходимости  $R$ -жадного алгоритма.

**Теорема 4.2.** Пусть  $X$  равномерно выпуклое, равномерно гладкое банахово пространство, для которого выполняется условие (20), и  $\mathcal{D}$  словарь в нем. Тогда для всех  $0 < t \leq 1$ ,  $q > 0$ , для любой  $f \in X$  и произвольной реализации  $R$ -жадного алгоритма выполняется

$$\lim_{m \rightarrow \infty} \|f - G_m^{RGA}(f, \mathcal{D}, t, q)\| = 0.$$

В параграфе 4.4. показывается, что условиям теоремы 4.2 удовлетворяет достаточно широкий класс банаховых пространств.

**Теорема 4.3.** Пространства  $l_p$ ,  $p > 2$ , обладают свойством (20).

Эти теоремы тесно связаны с результатами М. Ганичева и Н. Калтона [21], которые доказали справедливость утверждения теоремы 4.3 для всех  $p$ ,  $p > 1$ , и доказали сходимость  $D$ -ЖА. Подробности приведены в конце параграфа 4.1.

Хорошо известно, что  $X$ -жадный алгоритм сходится в любом конечномерном гладком банаховом пространстве (т.е. для всех  $f \in X$  и словарей  $\mathcal{D} \subset X$  выполняется  $\lim_{m \rightarrow \infty} \|f - G_m^{XGA}(f)\| = 0$ .)

В своей недавней работе [15] С. Диворф, Д. Куцарова, К. Шу-мац, В.Н. Темляков, П. Войташчек установили слабую сходимость  $X$ -жадного алгоритма для банаховых пространств, обладающих  $WN$ -свойством. С. Диворф, Е. Оделл, Т. Шлумпрехт, А. Зак [16] получили интересные результаты о сходимости  $X$ -жадного алгоритма в пространствах  $L_p(0, 1)$  для разреженных целевых функций. В тоже время



никаких общих “положительных” результатов о сходимости  $X$ -жадного алгоритма, в бесконечномерном банаховом пространстве, не являющимся гильбертовым, на сколько известно автору, получено не было. Более того, остается открытым вопрос о сходимости  $X$ -жадного алгоритма (для всех целевых функций) в пространстве  $L_p(0, 1)$ ,  $p \neq 2$ , по системе Хаара.

В параграфе 4.5 изучается скорость сходимости  $X$ -жадного алгоритма в  $L_p(0, 1)$ ,  $1 < p \leq 2$ , для системы Хаара и системы функций, пропорциональных индикаторам двоичных интервалов.

Напомним определение системы Хаара  $\mathcal{H}_p = \{H_n\}_{n=1}^\infty$  в пространстве  $L_p(0, 1)$ ,  $p > 1$ . Положим  $H_1(t) \equiv 1$ . Пусть  $m = 2^k + i$ ,  $k \geq 0$ ,  $1 \leq i \leq 2^k$ . Рассмотрим двоичные промежутки

$$\begin{aligned}\Delta_m &:= \Delta_k^i := \left[ \frac{i-1}{2^k}, \frac{i}{2^k} \right), \\ \Delta_m^+ &:= (\Delta_k^i)^+ := \left[ \frac{i-1}{2^k}, \frac{2i-1}{2^{k+1}} \right) = \Delta_{k+1}^{2i-1}, \\ \Delta_m^- &:= (\Delta_k^i)^- := \left[ \frac{2i-1}{2^{k+1}}, \frac{i}{2^k} \right) = \Delta_{k+1}^{2i}.\end{aligned}$$

Определим

$$H_m(t) = \begin{cases} 2^{k/p}, & t \in \Delta_m^+ \\ -2^{k/p}, & t \in \Delta_m^- \\ 0, & t \notin \Delta_m \end{cases}$$

Известно, что  $\overline{\text{span}} \mathcal{H}_p = L_p(0, 1)$ , поэтому, в силу равенств  $\|H_m\| = 1$ ,  $m \geq 1$ , система Хаара  $\mathcal{H}_p$  является словарем.

Для произвольного измеримого  $\Delta \subset [0, 1)$  обозначим через  $|\Delta|$  его меру Лебега. Определим

$$I_\Delta(t) = \begin{cases} |\Delta|^{-1/p}, & t \in \Delta \\ 0, & t \in [0, 1] \setminus \Delta. \end{cases}$$

Легко видеть, что  $\|I_\Delta\| = 1$ . В работе также изучается система функций  $\mathcal{I}_p$ , пропорциональных индикаторам двоичных интервалов. Положим

$$I_m := I_{\Delta_m}, \quad m \geq 2,$$

$$\mathcal{I}_p := \{I_m\}_{m=2}^\infty.$$

Приближение системой  $\mathcal{I}_p$  было рассмотрено Т.П. Лукашенко в статье [4].

Легко видеть, что для любого  $m \geq 1$

$$\Sigma_m(\mathcal{H}_p) \subset \Sigma_{2m}(\mathcal{I}_p) \quad (21)$$

(см. определение (6)), но в противоположную сторону аналогичные включения не верны

$$\Sigma_m(\mathcal{I}_p) \not\subset \Sigma_M(\mathcal{H}_p), \quad \forall M > m.$$

Для положительной функции  $\phi \in \mathcal{C}(\mathbb{R}_+)$  рассмотрим класс

$$\mathcal{A}_\phi(\mathcal{D}, X) = \mathcal{A}_\phi(\mathcal{D}) = \{f \in X : \sigma_m(f, \mathcal{D}, X) \leq \phi(m), m \geq 0\}.$$

**Теорема 4.4.** *Для любого  $p$ ,  $1 < p < \infty$ , существует такое  $\alpha_0 = \alpha_0(p) > 0$ , что для произвольной положительной убывающей функции  $\phi \in \mathcal{C}_1(\mathbb{R}_+)$ , удовлетворяющей неравенству*

$$\liminf_{t \rightarrow \infty} \frac{t\phi'(t)}{\phi(t)} = \liminf_{t \rightarrow \infty} t(\ln(\phi(t)))' \geq -\alpha_0, \quad (22)$$

для любой  $f \in \mathcal{A}_\phi(\mathcal{I}_p)$  и для произвольного  $m \geq 1$  будет справедлива оценка

$$\|f - G_m^{XGA}(f, \mathcal{I}_p)\| = \|f_m\| \leq C\phi(m)$$

где константа  $C > 0$  зависит от  $p$  и  $\phi(\cdot)$ , но не зависит от  $f$  и  $m$ .

Отметим, что для произвольных  $0 \leq \alpha < \alpha_0$ ,  $\beta \geq 0$ ,  $\alpha + \beta > 0$ , и  $C > 0$  для функции  $\phi \in \mathcal{C}_1(\mathbb{R}_+)$  вида

$$\phi(t) = C(t+1)^{-\alpha}(\ln(t+2))^{-\beta}$$

выполняется неравенство (22).

Из теоремы 4.4 выводится сходимость  $X$ -ЖА по системе  $\mathcal{I}_p$  для всех  $f \in L_p(0, 1)$ .

**Теорема 4.5.** *Для любого  $p$ ,  $1 < p < \infty$  и целевой функции  $f \in L_p(0, 1)$  имеет место равенство*

$$\lim_{m \rightarrow \infty} \|f - G_m^{XGA}(f, \mathcal{I}_p)\| = 0.$$

Имсет место следующий результат о скорости сходимости  $X$ -жадного алгоритма для системы Хаара в пространствах  $L_p(0, 1)$ ,  $1 < p \leq 2$ .

**Теорема 4.6.** Для любого  $p$ ,  $1 < p \leq 2$ , существует такое  $\alpha_0 = \alpha_0(p) > 0$ , что для произвольной положительной убывающей функции  $\phi \in C_1(\mathbb{R}_+)$ , удовлетворяющей неравенству

$$\liminf_{t \rightarrow \infty} \frac{t\phi'(t)}{\phi(t)} = \liminf_{t \rightarrow \infty} t(\ln(\phi(t)))' \geq -\alpha_0,$$

для любой  $f \in \mathcal{A}_\phi(\mathcal{I}_p, L_p(0, 1))$  и для произвольного  $m \geq 1$  будет справедлива оценка

$$\|f - G_m^{XGA}(f, \mathcal{H}_p)\| = \|f_m\| \leq \hat{c}_1 \phi(m) (\ln(m+1))^{\frac{p}{p-1}+2},$$

где константа  $\hat{c}_1 > 0$  зависит от  $p$  и  $\phi(\cdot)$ , но не зависит от  $f$  и  $m$ .

Из справедливости включения (21) вытекает

**Теорема 4.7.** Для любого  $p$ ,  $1 < p \leq 2$ , существует такое  $\alpha_0 = \alpha_0(p) > 0$ , что для произвольной положительной убывающей функции  $\phi \in C_1(\mathbb{R}_+)$ , удовлетворяющей неравенству

$$\liminf_{t \rightarrow \infty} \frac{t\phi'(t)}{\phi(t)} = \liminf_{t \rightarrow \infty} t(\ln(\phi(t)))' \geq -\alpha_0,$$

для любой  $f \in \mathcal{A}_\phi(\mathcal{H}_p, L_p(0, 1))$  и для произвольного  $m \geq 1$  будет справедлива оценка

$$\|f - G_m^{XGA}(f, \mathcal{H}_p)\| = \|f_m\| \leq \hat{c}_2 \phi(m) (\ln(m+1))^{\frac{p}{p-1}+2},$$

где константа  $\hat{c}_2 > 0$  зависит от  $p$  и  $\phi(\cdot)$ , но не зависит от  $f$  и  $m$ .

Из теоремы 4.7 вытекает сходимость  $X$ -жадного алгоритма на достаточно широком подклассе  $L_p(0, 1)$ ,  $1 < p \leq 2$ .

**Теорема 4.8.** Пусть задано  $p$ ,  $1 < p \leq 2$ , и функция  $f \in L_p(0, 1)$ , для которой существует такое  $\delta > 0$ , что

$$\sup_{m \geq 0} \sigma_m(f; \mathcal{H}_p) (\ln(m+2))^{\frac{p}{p-1}+2+\delta} < \infty.$$

Тогда

$$\lim_{m \rightarrow \infty} \|f - G_m^{XGA}(f, \mathcal{H}_p)\| = 0.$$

**Замечание 4.1.** При  $p = 2$  система Хаара  $\mathcal{H}_2$  является ортонормированным базисом в пространстве  $L_2(0, 1)$ , и хорошо известно, что для любой  $f \in L_2(0, 1)$

$$\|f - G_m^{XGA}(f, \mathcal{H}_2)\| = \|f - G_m^{PGA}(f, \mathcal{H}_2)\| = \sigma_m(f, H_2),$$

то есть теоремы 4.7 и 4.8 в случае  $p = 2$  не являются оптимальными.

**Замечание 4.2.** Если в банаховом пространстве  $X$  задан базис Шаудера  $\mathcal{B} = \{\psi_k\}_{k=1}^\infty$ , то  $m$ -членные приближения к  $f \in X$  по системе  $\mathcal{B}$ , можно получать с помощью метода, основанного на разложении  $f$  в ряд по базису  $\mathcal{B}$

$$f = \sum_{k=1}^{\infty} c_k(f) \psi_k,$$

и существенно отличного от изложенного выше  $X$ -жадного алгоритма. Этот метод основан на перепорядочении коэффициентов разложения в порядке убывания модуля

$$|c_{k_1}(f)| \geq |c_{k_2}(f)| \geq \dots$$

В качестве  $m$ -членного приближения  $G_m^*(f, \mathcal{B})$  берется

$$G_m^*(f, \mathcal{B}) = \sum_{j=1}^m |c_{k_j}(f)| \psi_{k_j}.$$

В.Н. Темляков [28] доказал, что для произвольного  $p > 1$  найдется такое число  $C(p) > 0$ , что для всех  $f \in L_p(0, 1)$  и  $m \geq 1$  выполняется неравенство

$$\|f - G_m^*(f, \mathcal{H}_p)\| \leq C(p) \sigma_m(f, \mathcal{H}_p).$$

В параграфе 4.6 излагается схема доказательства теоремы 4.6. Указывается, что ключевую роль в изучении сходимости  $X$ -жадного алгоритма по системе Хаара играет следующий результат об  $m$ -членных линейных комбинациях функций из системы  $\mathcal{I}_p$ .

**Теорема 4.9.** Для любого  $p$ ,  $1 < p \leq 2$ , существуют такие числа  $\hat{c}_3, \hat{c}_4 > 0$ , что для любого  $n \in \mathbb{N}$  и  $h \in \Sigma_n(\mathcal{I}_p)$ ,  $\|h\| = 1$  найдется

конечное множество индексов  $\Lambda_0 \subset \mathbb{N}$ , для которого выполняются следующие неравенства

$$r(h, H_\lambda) \geq \frac{\widehat{c}_3}{n}, \text{ для любого } \lambda \in \Lambda_0,$$

$$\sum_{\lambda \in \Lambda_0} r(h, H_\lambda) \geq \widehat{c}_4 (\ln(n+1))^{-\frac{p}{p-1}-1},$$

Этот результат доказывается в параграфах 4.7, 4.8. С его помощью в параграфе 4.9 доказывается теорема

**Теорема 4.10.** Для любого  $p$ ,  $1 < p \leq 2$ , существуют такие  $\widehat{c}_5, \widehat{c}_6 > 0$ , что для любого  $n \in \mathbb{N}$ ,  $h \in \Sigma_n(\mathcal{I}_p)$ ,  $\|h\| = 1$  и функции  $f \in L_p(0, 1)$  такой, что

$$\|f - h\| < \widehat{c}_5 (\ln(n+1))^{-\frac{p}{p-1}-2},$$

найдется  $\lambda_0 \in \mathbb{N}$ , для которого

$$r(f, H_{\lambda_0}) \geq \frac{\widehat{c}_6}{n}.$$

В параграфе 4.10 теорема 4.10 применяется для доказательства теоремы 4.6, с помощью которой там же устанавливается справедливость теоремы 4.8.

В параграфе 4.11 приводится доказательство теорем 4.4 и 4.5.

# Литература

- [1] Галатенко В. В. “Об орторекурсивном разложении по некоторой системе функций с ошибками при вычислении коэффициентов”. // *Матем. сб.* 2004. Т. 195:7. С. 21–36.
- [2] Исмагилов Р.С. “Поперечники множеств в линейных нормированных пространствах и приближение функций тригонометрическими многочленами” // *УМН* 1974. Т. 29. № 3. С. 161–178.
- [3] Кашин В.С. “Об аппроксимационных свойствах полных ортонормированных систем.” // *Тр. МИАН СССР*. 1985. Т. 172. С. 187–191.
- [4] Лукашенко Т. П. “ О свойствах орторекурсивных разложений по неортогональным системам” // *Вестник Моск. Ун-та. Сер. I. Матем. Механ.* 2001. № 1. С. 6–10.
- [5] Майоров В.Е. “ Тригонометрические поперечники соболевских классов  $W_p^r$  в пространстве  $L_q$ .” // *Матем. заметки* 1986. Т. 40. № 2. С. 161–173.
- [6] Осколков К.И. “ Аппроксимативные свойства суммируемых функций на множествах полной меры.” // *Матем. сб.* 1977. Т. 103(145):4(8). С. 563–589.
- [7] Сильниченко А. “О скорости сходимости жадных алгоритмов” // *Математические заметки* 2004. Т. 76. № 4. С. 628–631.
- [8] Стечкин В.С., Стечкин С.Б. “Среднее квадратическое и среднее арифметическое” // *Докл. АН СССР* 1961. Т. 137. № 2. С. 287–290.
- [9] Стечкин С.Б. “Об абсолютной сходимости ортогональных рядов” // *Докл. АН СССР* 1955. Т. 102. № 1. С. 37–40.

- [10] Barron A. "Universal approximation bounds for superposition of  $n$  sigmoidal functions" // *IEEE Transactions on Information Theory*. 1993. V. 39. P. 930–945.
- [11] Barron A. "Approximation and estimation bounds for artificial neural networks" // in *Proc. Fourth Annual Workshop on Computational Learning Theory*. M. Kaufmann. 1991. P. 243–249.
- [12] Barron A., Cohen A., Dahmen W., DeVore R. "Approximation and learning by greedy algorithms" // *Annals of Statistics* 2008. V. 36. P. 64–94.
- [13] Bergh J., Löfström J. "Interpolation spaces" // Springer Verlag. Berlin. 1976.
- [14] DeVore R. A., Temlyakov V. N. "Some remarks on Greedy Algorithms" // *Advances in Computational Mathematics*. 1996. V. 5. P. 173–187.
- [15] Dilworth S., Kutzarova D., Shuman K., Temlyakov V. N., Wojtaszchuk P. "Weak Convergence of Greedy Algorithms in Banach spaces" // *Journal of Fourier Analysis and Applications*. 2008. V. 14. P. 608–629.
- [16] Dilworth S., Odell E., Schlumprecht Th., Zsák A. "On the Convergence of Greedy Algorithms for Initial Segments of the Haar Basis" // *Math. Proc. Cambridge Philos. Soc.* accepted 2009.
- [17] Donoho D. L., Elad M., Temlyakov V. N. "Stable recovery of sparse overcomplete representations in the presence of noise" // *IEEE Trans. Inform. Th.* 2006. V. 52:1. P. 6–18.
- [18] Donoho D. L., Elad M., Temlyakov V. N. "On Lebesgue-type inequalities for greedy approximation" // *Journal of Approximation Theory* 2007. V. 147:2. P. 185–195.
- [19] V. V. Dubinin "Greedy Algorithms and Applications". // *Ph.D. thesis, University of South Carolina*. 1997.
- [20] Friedman J. H., Stuetzle W. "Projection pursuit regression" // *J. Amer. Statist. Assoc.* 1981. V. 76. P. 817–823.

- [21] Ganichev M., Kalton N.J. "Convergence of the weak dual greedy algorithm in  $L_p$ -spaces" // *Journal of Approximation Theory* 2003. V. 124. P. 89–95.
- [22] Gilbert A.C., Muthukrishnan M., Strauss J. "Approximation of functions over redundant dictionaries using coherence" // *Proc. 14th Annu. ACM-SIAM Symp. Discrete Algorithms*. 2003. P. 243–252.
- [23] Gribonval R., Nielsen M. "On the strong uniqueness of highly sparse expansions from redundant dictionaries." // *Proc. Int Conf. Independent Component Anal. (ICA'04)* 2004.
- [24] Jones L., "On a conjecture of Huber concerning the convergence of projection pursuit regression" // *Ann. Statist.* 1987. V. 15. P. 880–882.
- [25] Konyagin S. V., Temlyakov V. N. "Rate of convergence of Pure Greedy Algorithm" // *East J. Approx.* 1999. V. 5. P. 493–499.
- [26] Mallat S., Zhang Z. "Matching pursuit with time-frequency dictionaries" // *IEEE Trans. Signal Process.* 1993. V. 41. № 12. P. 3397–3415.
- [27] Schmidt E. "Zur Theorie der linearen und nichtlinearen Integralgleichungen. I." // *Math. Ann.* (1906-1907). 63. P. 433–476.
- [28] Temlyakov V.N. "The best  $m$ -term approximation and greedy algorithms" // *Advances in Comp. Math.* 1998. V. 8. P. 249–265.
- [29] Temlyakov V.N. "Greedy approximation" // *Acta Numerica*. 2008. V. 17. P. 235–409.
- [30] Temlyakov V.N. "Greedy expansions in Banach Spaces" // *IMI-preprint*. 2003. No. 6. P. 1–21.
- [31] Temlyakov V.N., Zheltov P. "On performance of greedy algorithms" <http://dsp.rice.edu/cs> //submitted to *Journal of Approximation Theory*. 2010
- [32] Tropp J.A. "Greed is good: algorithmic results for sparse approximation" // *IEEE Trans. Inform. Th.* 2004. V. 50:10. P. 2231–2242.



- [33] Лившиц Е. Д., Темляков В. Н. “ О сходимости слабого гриди-алгоритма” // *Труды математического института РАН им. В.А. Стеклова* . 2001. Т. 232. С. 236–247.
- [34] Лившиц Е. Д. “ О сходимости X-гриди-алгоритма по системе сплайнов” // *Труды XXIV конференции молодых ученых мех-мат факультета МГУ им М.В. Ломоносова* . 2002. С. 113–116.
- [35] Galatenko V.V., Livshitz E.D. “On the convergence of approximative weak greedy algorithms” // *East J. Approx.* 2003. V. 9. No. 1. P. 43–49.
- [36] Livshitz E. D., Temlyakov V. N., “Two lower estimates in greedy approximation” // *Constructive Approximation*. 2003. V. 19. P. 509–524.
- [37] Галатенко В. В., Лившиц Е. Д. “Обобщенные приближенные слабые жадные алгоритмы” // *Математические заметки*. 2005. Т. 78. № 2. С. 186–201.
- [38] Галатенко В. В., Лившиц Е. Д., Шмырев Н. В. “Сходимость обобщенных приближенных слабых жадных разложений по ортонормированным системам.” // *Технологии программирования, учетные системы, ортонормированные системы*. 2008. НИИ системных исследований РАН. С. 149–163.
- [39] Livshits E.D. “On the optimality of Orthogonal Greedy Algorithm for M-coherent dictionaries ” //2010. <http://arxiv.org/abs/1003.5349>
- Основные публикации по теме диссертации (из перечня ВАК).**
- [40] Лившиц Е. Д. “О сходимости гриди-алгоритмов в банаховых пространствах” // *Матем. заметки* 2003. Т. 73. № 3. С. 371–389.
- [41] Лившиц Е. Д. “О скорости сходимости чисто жадного алгоритма” // *Матем. заметки* 2004. Т. 76. № 4. С. 539–552.
- [42] Лившиц Е. Д. “О возвратном жадном алгоритме” // *Изв. РАН. Серия матем.* 2006. Т. 70. № 1. С. 95–116.

- [43] Лившиц Е.Д. “Об оптимальности жадного алгоритма для некоторых классов функций” // *Матем. сборник* 2007. Т. 198. № 5. С. 95–114.
- [44] Лившиц Е.Д. “Об  $n$ -членном приближении с неотрицательными коэффициентами.” // *Матем. заметки* 2007. Т. 82. № 3. С. 373–382.
- [45] Лившиц Е.Д. “О жадного алгоритме в пространстве  $L_p[0, 1]$ .” // *Матем. заметки* 2009. Т. 85. № 5. С. 788–791.
- [46] Лившиц Е.Д. “О нижних оценках скорости сходимости жадных алгоритмов” // *Изв. РАН. Серия матем.* 2009. Т. 73. № 6. С. 125–144.
- [47] Лившиц Е.Д. “О сходимости жадного алгоритма по системе Хаара в пространствах  $L_p(0, 1)$ ” // *Матем. сборник* 2010. Т. 201. № 2. С. 99–130.
- [48] Лившиц Е.Д. “О жадном алгоритме для словарей с ограниченной совокупной когерентностью” // *Матем. заметки* 2010. Т. 87. № 5. С. 792–795.
- [49] Лившиц Е.Д. “Реализуемость жадных алгоритмов” // *Труды ИММ УрО РАН*. 2010. Т. 16:4. С. 228–236.

Подписано в печать 09.03.11  
Объем: 1,5 усл.печ.л.  
Тираж: 100 экз. Заказ № 765  
Отпечатано в типографии «Реглет»  
119526, г. Москва, пр-т Вернадского, 39  
(495) 363-78-90; [www.reglet.ru](http://www.reglet.ru)